

A REVIEW OF THE TURKISH SCALES ON ARTIFICIAL INTELLIGENCE IN EDUCATION

Dr. Derya UYGUN
Ministry of National Education, Eskisehir, Turkey
derya.ul@gmail.com

ABSTRACT

Background: The increasing integration of artificial intelligence in education highlights the pressing need for effective policies and practices. However, much of the existing research has focused mainly on international contexts, which has created a gap concerning Turkish scales. Therefore, reviewing Turkish scales is essential for designing solutions specific to the unique characteristics of Turkey's education system.

Methods: This study aims to systematically review and evaluate scales in Turkish that measure various dimensions of artificial intelligence in education. The study employed a systematic review methodology, searching various databases to identify relevant scales. A total of 16 scales were included. Each scale was examined in detail to assess its structure, target constructs, and psychometric robustness.

Results: The scales reviewed primarily measured constructs such as artificial intelligence awareness, anxiety, literacy, attitudes, and readiness, structured across 2 to 7 factors. Target populations included teachers, university students, pre-service teachers, and adults, with sample sizes ranging from 85 to 653 participants. Reliability and validity analyses demonstrated that these scales are robust and effective tools for measuring AI-related dimensions in education.

Conclusion: This review highlights a predominant focus on adapting existing scales to the Turkish context, reflecting a need for more original scale development tailored to specific educational requirements. These findings provide valuable insights for future research aimed at enhancing artificial intelligence in education.

Keywords: Artificial intelligence, Artificial Intelligence in Education, Turkish Scales, Systematic Review, AI Literacy, Scale Adaptation, Scale Development

INTRODUCTION

The integration of artificial intelligence (AI) into education has emerged as a transformative force, reshaping pedagogical practices and learning experiences across various educational contexts. The Turkish education system, like many others globally, is witnessing a paradigm shift as educators and institutions increasingly recognize the importance of artificial intelligence in education (AIEd). This shift is crucial not only for enhancing teaching practices but also for preparing students for a future where AI will play a central role in various professions (Arvin, 2023; Uygun, 2024).

The rapid advancement of AIEd necessitates a comprehensive assessment of both individual and institutional competencies within this domain. However, the scales currently employed to evaluate and measure the impact of AI in educational contexts are predominantly developed in international settings and primarily designed in English (Çelebi et al., 2023). This reliance on such scales may result in a misalignment with the specific nuances and requirements of the Turkish education system. Moreover, the scarcity of Turkish-language scales significantly hampers the ability to gauge the awareness, literacy, attitudes, and readiness levels of teachers, students, and educational administrators concerning AI (Polatgil & Güler, 2023; Erdoğan & Ekşioğlu, 2024).

In the endeavor to adapt or develop Turkish scales pertinent to AIEd, it has become evident that the existing scales have not undergone systematic examination regarding their qualities and scope. The absence of a thorough evaluation presents a considerable challenge for researchers and educators, thereby creating a gap in the guidance necessary for selecting and developing effective measurement tools. This assertion is supported by the work of Karademir et al. (2022), who emphasize the significance of evaluating educational tools to ensure they meet the needs of the educational environment effectively. They argue that the selection of appropriate instruments is crucial for making informed decisions aimed at improving educational quality. This perspective underscores the necessity for systematic evaluations of existing scales, particularly in the context of adapting or developing measurement tools for specific educational settings.

This study seeks to bridge this gap by systematically reviewing Turkish scales related to AIEd. By analyzing their characteristics, psychometric properties, and target constructs, this research aims to provide a comprehensive overview of the current state of measurement tools in AIEd. Ultimately, the findings will guide researchers, educators, and policymakers in selecting and developing robust instruments that can effectively support the integration of AI into Turkish educational practices.

Literature Review

In the context of education, AI supports individualized learning and the identification of learning difficulties, offers personalized educational programs such as materials and activities tailored to learning styles, and provides an accessible education model during extraordinary situations like pandemics and disasters (Saçan et al., 2022). A review of the literature on the subject reveals a noticeable increase in research on AIED in recent years (Arik & Seferoğlu, 2022; Wang et al., 2024). Oruç et al. (2024) examined 51 Turkey-based studies conducted between 2020 and 2024, revealing an increase in academic research on AIED. Of these studies, 60.8% were published in Turkish, and 13.7% were published in English. Most of the research was conducted on teachers (22.5%) and university students (10%), with purposive sampling being the most used approach. Most of the studies focused on opinion determination (19.2%), systematic reviews (15.4%), and scale development (9%), highlighting the growing interest in creating tools to measure various dimensions of AIED. This aligns with the current study's objective to systematically review and evaluate Turkish scales developed or adapted for assessing AI-related competencies and attitudes in the educational context.

Similarly, Lintner's (2024) systematic review underscores the importance of evaluating and enhancing the quality of AI literacy scales, further emphasizing the need for robust and contextually relevant tools to effectively assess AI-related competencies, which aligns with the objectives of this study in the Turkish educational context. The systematic review conducted by Lintner (2024) examines the scales used to assess AI literacy and evaluates their quality using the COSMIN tool. The study analyzes 22 research articles validating a total of 16 scales designed for the public, higher education, and secondary school students, as well as teachers. Most scales demonstrated positive results in terms of structural validity and internal consistency. However, critical features such as content validity, reliability, structural validity, and responsiveness were rarely tested, while cross-cultural validity and measurement error were not tested at all. Notably, most of the scales were self-report-based, with only a few performance-based scales available.

As correctly noted, Lintner's (2024) study focuses exclusively on AI literacy and primarily examines scales developed for diverse populations in general contexts. In contrast, this review adopts a more specific focus, aiming to analyze Turkish scales developed or adapted specifically for AIED and to summarize their unique characteristics. This review focuses on the adaptation and development of AI scales in education specifically tailored for the Turkish educational landscape.

A review of the literature reveals that there is no existing study summarizing Turkish scales related to AIED. However, such a review could provide valuable insights into the scope and characteristics of existing scales, serving as a guide for researchers interested in this field. Researchers could refer to this review to select the most appropriate scale for their studies and gain detailed information about the target audience, psychometric properties, and application context of the scales. This would enable them to conduct accurate and effective measurements while saving time and resources. Additionally, such a review would not only facilitate the use of existing scales but also serve as a foundation for the development of new scales. It would contribute significantly to the effective evaluation of AIED within the Turkish context, thereby enriching the scientific knowledge in this field. Therefore, the aim of this study is to systematically review and evaluate Turkish scales developed or adapted for measuring various dimensions of AIED, providing insights into their characteristics, psychometric properties, and application contexts.

METHOD

Research Model

In this study, a descriptive content analysis method was employed to systematically review and evaluate Turkish scales related to AIED. Descriptive content analysis is a systematic research method that involves examining studies on a specific topic to evaluate their trends, characteristics, and findings in a descriptive manner. This method was chosen to gain an in-depth understanding of existing Turkish scales, identify their similarities and differences, and provide a comprehensive overview of their scope and psychometric properties (Yeşilyurt, 2021). Instead of sampling, the strategy of reaching all available studies within the research universe was adopted.

Data Collection

In this study, the criterion sampling method, one of the purposive sampling techniques, was used to select the studies to be included. The criterion sampling method aims to select items that meet predetermined criteria set by the researcher (Patton, 2014). The literature review was conducted using the databases Google Scholar, YÖKTEZ, Tr Dizin, and Dergipark with the following keywords:

“Artificial intelligence education scale in Turkish”

“Turkish version artificial intelligence in education survey”

“Turkish adaptation artificial intelligence in education questionnaire”

The criteria determined by the researchers for selecting the studies to be included in this review are as follows:

- Studies available in the Google Scholar, YÖKTEZ, Tr Dizin, and Dergipark databases.
- Studies identified using the keywords.
- Research focused on AIED.
- Scales developed or adapted in Turkish.
- Studies with full-text accessibility.

The database searches were performed in April 2025, covering studies published up to that date. Only studies meeting all these criteria were included in the scope of this research. Figure 1 PRISMA flowchart shows the study identification procedure.

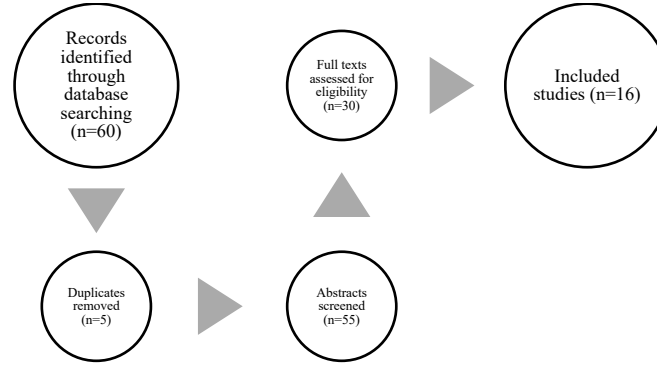


Figure 1. PRISMA Flowchart

A total of 16 articles meeting the specified inclusion criteria were included in the study. The studies included in the research are presented in Table 1.

Table 1. Articles Included in the Study

No	Codes	Reference	The name of the scale
1	A1	Ferikoğlu and Akgün, (2022)	AI Awareness of Teachers
2	A2	Dülger and Köklü, (2023)	Opinions of School Administrators and Teachers on the Use of AIED
3	A3	Özüdoğru and Yıldız Durak (2024)	AI Readiness Scale
4	A4	Karaoğlu Yılmaz and Durak (2024)	Generative AI Acceptance Scale
5	A5	Terzi (2020)	AI Anxiety Scale
6	A6	Akkaya et al. (2021)	AI Anxiety Scale
7	A7	Kaya et al. (2024)	Threats of AI Scale
8	A8	Hopcan et al. (2021)	Machine Learning Attitude Scale
9	A9	Gökçearslan et al. (2024)	GenAI Literacy Scale
10	A10	Çelebi et al. (2023)	AI Literacy Scale
11	A11	Polatgil, M., Güler, A. (2023)	AI Literacy Scale
12	A12	Enis Erdoğan and Ekşioğlu (2024)	AI Literacy Scale
13	A13	Karaoğlu Yılmaz and Yılmaz (2023)	AI Literacy Scale
14	A14	Bilici et al. (2023)	AI-TPACK Scale
15	A15	Alan et al. (2024)	Attitudes Toward AI Scale (ATAI)
16	A16	Kaya et al. (2022)	General Attitudes toward AI Scale (GAAIS)

Data Analysis

A research classification chart developed by the researchers was used as a data collection tool in this study. The identified studies were classified and analyzed in detail by recording them into the classification chart based on the following categories: year and type of study, publication language, published journal/university, research purpose, research approach, research method, sample group, sampling method, sample size, data collection tool, analysis method, and validity-reliability measures. The data obtained were subjected to a detailed content analysis, and the findings were presented in the form of figures and tables.

Validity and Reliability

To ensure the validity and reliability of the study, multiple strategies were implemented. The classification chart was reviewed by two experts with experience in qualitative research and AIED to ensure the accuracy and consistency of the data entries. Additionally, coding consistency was cross-checked by independent researchers, and any discrepancies were resolved through discussion. Detailed documentation of the analysis process was maintained to enhance transparency. This systematic review involved the analysis of existing literature and did not include any human participants or personal data collection. Therefore, ethical approval was not required in accordance with institutional and national research ethics guidelines.

Table 2. Summary of Scale Characteristics

Description of the scales			Participants				Factors	
Scale	Codes	Type	Items	Target Population	N	Age	No of fac	Factor description
AI Awareness of Teachers	A1	Developed Self-report five-point Likert scale.	51	Teachers	561	22-64	4	Theoretical knowledge, practical knowledge, beliefs-attitudes, and integration ability
Opinions of School Administrators and Teachers on the Use of AIEd AI Readiness Scale	A2	Developed Self-report five-point Likert scale.	28	School administrators and teachers	225	Not Reported	4	Positive aspects of AI, harms of AI, scope of AI, concept of AI
	A3	Adaptation Self-report five-point Likert scale.	18	Pre-service teachers	Not Reported	18- 25	4	Cognition, Ability, Vision, and Ethics in Teaching
Generative AI Acceptance Scale	A4	Developed Self-report five-point Likert scale	20	University students	627	Not Reported	4	Effort Expectancy Facilitating Conditions Performance Expectancy Social Influence
AI Anxiety Scale	A5	Adaptation Self-report seven-point Likert scale	21	Teachers	222	22-57	4	Learning, Job Replacement, Sociotechnical Blindness, AI Configuration
AI Anxiety Scale	A6	Adaptation Self-report	21	Teachers	490	Not Reported	4	Learning, Job Replacement, Sociotechnical Blindness, AI Configuration
Threats of AI Scale	A7	Adaptation Self-report five-point Likert scale	12	Adults	285	18-45	4	Identification Prediction Recommendation Decision Making
Machine Learning Attitude Scale	A8	Adaptation Self-report five-point Likert scale	23	University students	309	Not Reported	6	Interest in Technology, Gender Role of Technology, Importance and Impact of Technology, Courses Technology-related Career Paths Technology and Creative Activities
Gen AI Literacy Scale	A9	Adaptation Self-report	12	University students	297	21	4	Awareness, Use, Evaluation, Ethics
AI Literacy Scale	A10	Adaptation Self-report	12	Non-expert adults	402	Not Reported	4	Awareness, Use, Evaluation, Ethics
AI Literacy Scale	A11	Adaptation Self-report	12	Adults	536	18-60	4	Awareness, Use, Evaluation, Ethics
AI Literacy Scale	A12	Adaptation Self-report	12	Teachers	226	Not Reported	4	Awareness, Use, Evaluation, Ethics
AI Literacy Scale	A13	Adaptation Self-report	31	High school, University students, University graduates	653	14-54	3	Technical Understanding, Critical Evaluation, Practical Application
AI-TPACK Scale	A14	Adaptation Self-report five-point Likert scale	39	STEM Teachers	85	Not Reported	7	Knowledge Components related to AI in Education
Attitudes Toward AI Scale (ATAI)	A15	Developed	36	Teachers	576	Not Reported	4	AI in Daily Life, Education, Resistance, Awareness
General Attitudes toward AI Scale (GAAIS)	A16	Adaptation Self-report	20	University students	350	18-51	2	Positive and Negative Attitudes Toward AI

FINDINGS

The Table 2 summarizes 16 scales related to AIEd. Among these, 12 scales are adaptations (A3, A5, A6, A7, A8, A9, A10, A11, A12, A13, A14, A16), while 4 (A1, A2, A4, A15) are original developments.

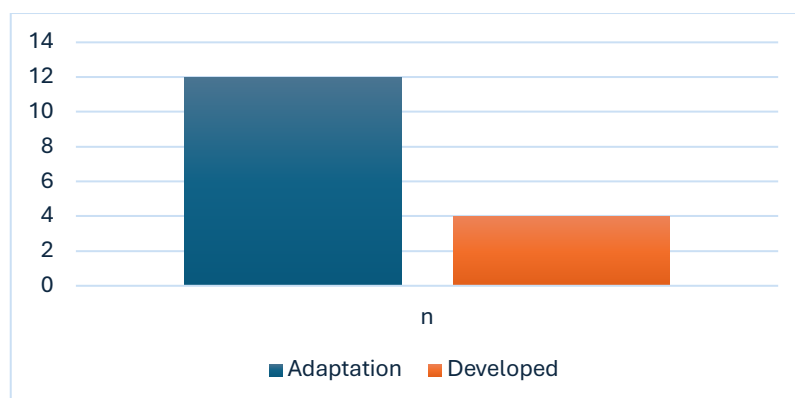


Figure 2. Types of the scales

As illustrated in Figure 2, all the scales employed in the reviewed studies are self-report instruments. Among these, 12 scales are adaptations of existing measures (adapted self-report), while 4 scales were newly developed for the specific research context (developed self-report). These instruments rely exclusively on participants' self-reported data. Although self-report measures offer practical advantages for assessing perceptions and attitudes, they are inherently susceptible to response biases, particularly social desirability bias, which may affect the validity of the findings.

When examining Table 2, the number of items in the scales range from 12 to 51. The longest scale is A1, comprising 51 items, while the shortest scales—A9, A10, A11, and A12—each consist of 12 items. On average, most of the scales are structured with approximately 20 items.

All scales reported the number of factors extracted, with the majority comprising two to four factors. The psychometric analyses conducted suggest that structural validity was adequately established across the instruments. The most frequently identified factor themes include awareness, anxiety, literacy, attitudes, and readiness. Additionally, some scales incorporated other thematic dimensions such as Configuration, Job Loss, Gender Roles, and Social Impact.

As seen in Figure 3, the most frequently targeted groups in the reviewed studies were teachers (A1, A2, A5, A6, A12, A15, A14), university students (e.g., A4, A8, A9, A13, A16), and adults/general population (e.g., A7, A11, A12). Among these, teachers and university students were the most examined populations. Notably, younger age groups such as primary and secondary school students have not been the primary focus in these studies. The number of participants across the studies ranged from 28 to 653. The largest sample was reported in study A13, which included 653 participants, while the smallest sample was in study A2, with 28 participants consisting of school administrators and teachers. Overall, the sample sizes were generally adequate to support psychometric analyses.

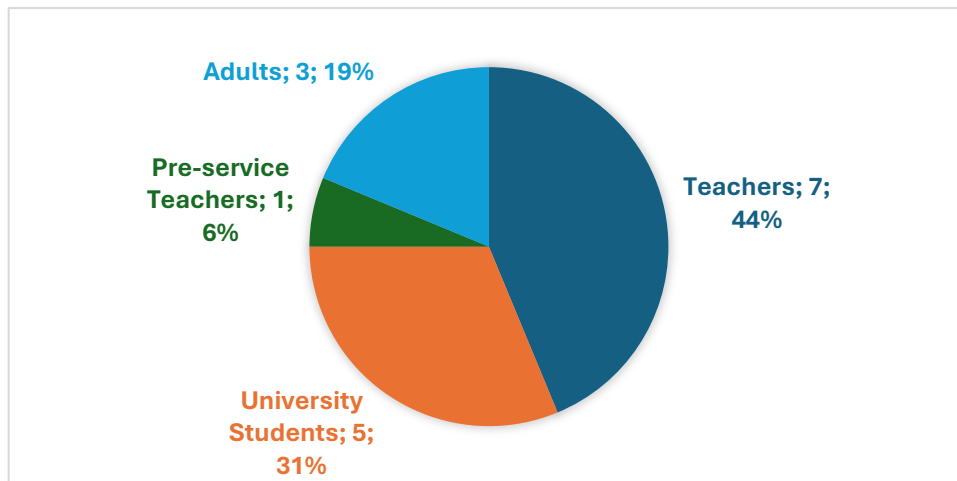


Figure 3. Target population

The Table 2 provides detailed information about AI-related scales in education, including their types, number of items, target populations, age ranges, and measured factor. In addition, Table 3 presents a summary of the psychometric properties of the reviewed scales, including internal consistency coefficients, factor structures, and model fit indices, to offer a concise overview of their methodological rigor.

Table 3. Psychometric Properties of the Reviewed Scales

Codes	The name of the scale	Cronbach's Alpha (α)	Fit indices (CFI, RMSEA)
A1	Teachers' AI Awareness Scale	0.986	$\chi^2/df=3.67$, CFI=0.903, RMSEA=0.069
A2	Opinion Scale for AI in Education	0.922	KMO=0.913, GFI=0.862, CFI=0.915, RMSEA=0.073
A3	AI Readiness Scale	0.967	RMSEA=0.08, CFI=0.99, NFI=0.98
A4	AI Acceptance Scale	0.97	RMSEA=0.08, CFI=0.99, NFI=0.98
A5-A6	AI Anxiety Scales	A5: 0.96, A6: 0.937	A5: CFI=0.94, RMSEA=0.084; A6: CFI=0.95, RMSEA=0.078
A7	Threats of AI Scale	0.74-0.83	RMSEA=0.041, CFI=0.998
A8	Machine Learning Attitude Scale	0.76-0.89	RMSEA=0.07, CFI=0.99
A9-A10-A11-A12-A13	AI Literacy Scales	A9: 0.74, A10: 0.85, A11: 0.939, A12: 0.861, A13: 0.97-0.99	A11: RMSEA=0.041, CFI=0.998 (best fit)
A14	AI-TPACK Scale	0.98	verified the seven-factor structure
A15- A16	AI Attitudes Scales	A15: 0.962, A16: 0.88	A15 is more comprehensive

As presented in Table 3, most of the reviewed scales demonstrate high internal consistency, typically with Cronbach's Alpha values above .85, and strong model fit indices (e.g., CFI > .95, RMSEA < .08). The most common factor structure consists of four dimensions, indicating a multidimensional approach to assessing AI-related constructs in education. Among the AI literacy scales, five distinct adaptations were identified. Notably, A13 (Karaoğlu Yılmaz & Yılmaz, 2023) stands out with the highest reliability ($\alpha = 0.97-0.99$) and the largest sample size ($N = 653$), making it the most robust in this category. In terms of teacher-oriented instruments, both A1 (AI Awareness Scale) and A14 (AI-TPACK Scale) exhibit excellent internal consistency ($\alpha > 0.98$) and solid construct validity, confirming their applicability in educational settings. Regarding the anxiety-related scales, although both A5 (Terzi, 2020) and A6 (Akkaya et al., 2021) are psychometrically sound, A6 reports slightly better model fit indices, suggesting it may be a more precise tool for measuring AI-related anxiety. Overall, the psychometric data in Table 3 provide a comparative foundation for selecting appropriate instruments based on reliability, validity. Following the table, the scales are further discussed in terms of their validity and reliability values, providing a comparative analysis of their robustness and applicability.

The Teachers' Artificial Intelligence Awareness Scale (A1) was developed to measure teachers' levels of awareness about AI. The scale was created by Deren Ferikoğlu and Ergün Akgün (2022). Its reliability was confirmed with a Cronbach's Alpha coefficient of 0.986. In the study conducted by Ferikoğlu & Akgün (2022), in determining the fit of the model, Comparative Fit Index (CFI), and Root Mean Square Error of Approximation

(RMSEA) were utilized. χ^2/df ratio was 3,67, which indicates a good fit between the model and the data. The CFI value was calculated as .903 and RMSEA was .069. The results indicated that the model demonstrated a high level of fit. The findings of the study conducted by Ferikoğlu & Akgün (2022) demonstrate the validity of the scale and the appropriateness of the model in fitting the data.

The Opinion Scale for the Use of Artificial Intelligence in Education (A2) was developed to evaluate the perspectives of school administrators and teachers regarding the use of AIED contexts. Designed by Eda Demir Dülger and Muharrem Köklü (2023), the scale has demonstrated high reliability with a Cronbach's Alpha coefficient of 0.922. Its validity was confirmed through Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), yielding a KMO value of 0.913 and a total explained variance of 64%. Fit indices from CFA, including Goodness of Fit Index (GFI) = 0.862, CFI = 0.915, and RMSEA = 0.073, support the scale's robustness. These results indicate that the scale is a valid and reliable tool for assessing opinions on the use of AIED.

AI Readiness Scale (A3) for Preservice Teachers was developed to measure the readiness of preservice teachers to use AI technologies in education. Originally created by Wang et al. (2023), the scale was adapted to Turkish by Özüdoğru and Yıldız Durak (2023). Its reliability was confirmed with a Cronbach's Alpha coefficient of 0.967 (overall scale), and sub-dimensions ranged from 0.819 to 0.933. Construct validity was demonstrated through CFA with acceptable fit indices: RMSEA = 0.08, CFI = 0.99, and Normed Fit Index (NFI) = 0.98. These results indicate that the scale is a valid and reliable tool for assessing the AI readiness of preservice teachers.

The Generative Artificial Intelligence Acceptance Scale (A4) was developed to measure students' acceptance of generative AI applications. Created by Karaoğlu Yılmaz et al. (2023), the scale is grounded in the Unified Theory of Acceptance and Use of Technology (UTAUT). Its reliability was confirmed with a Cronbach's Alpha coefficient of 0.97, while test-retest reliability yielded a coefficient of 0.95. Construct validity was demonstrated through EFA, which accounted for 78.349% of total variance, and CFA, which showed excellent fit indices: RMSEA = 0.08, CFI = 0.99, and NFI = 0.98. These results indicate that the scale is a valid and reliable tool for assessing the acceptance of generative AI applications.

Both **Artificial Intelligence Anxiety Scales (A5-A6) (AIAS)** by Wang and Wang (2019) adapted into Turkish by Terzi (2020) (A5) and by Akkaya et al. (2021) (A6) aim to measure anxiety levels related to AI. The AIAS by Terzi has a total Cronbach's Alpha reliability of 0.96, with fit indices indicating good model compatibility (e.g., CFI = 0.94, RMSEA = 0.084). Similarly, the AIAS by Akkaya et al. (2021) also demonstrates strong reliability, with a Cronbach's Alpha of 0.937 and fit indices (e.g., CFI = 0.950, RMSEA = 0.078) within acceptable ranges. Despite both scales being robust, Akkaya et al.'s adaptation (AIA) shows slightly better construct validity through more stringent fit indices and higher sub-dimension reliability. Therefore, the AIA Scale by Akkaya et al. (2021) is recommended as a slightly more reliable and valid tool for measuring AI-related anxiety.

Threats of Artificial Intelligence Scale (A7), originally developed by Kieslich et al. (2021), adapted into Turkish by Kaya et al. (2024), aims to measure individuals' fear levels toward AI applications. The reliability analysis of the Turkish adaptation showed high consistency, with Cronbach's Alpha values ranging from 0.74 to 0.83 for the sub-dimensions. CFA results indicated an excellent model fit, with indices such as RMSEA = 0.041 and CFI = 0.998. These findings confirm that the scale is both valid and reliable for use in evaluating AI-related fears in Turkish culture.

The Machine Learning Attitude Scale (A8), originally developed by Lee (2019), was adapted into Turkish by Hopcan, Polat, and Türkmen (2021) to measure university students' attitudes toward machine learning technologies. The scale's reliability was confirmed with Cronbach's Alpha coefficients ranging from 0.76 to 0.89 for its sub-dimensions, and CFA results indicated excellent model fit with indices such as RMSEA = 0.07 and CFI = 0.99. These findings demonstrate that the scale is a valid and reliable tool for assessing attitudes toward machine learning technologies.

Five different studies adapted the **Artificial Intelligence Literacy Scale (A9, A10, A11, A12, A13)** by Wang et al. into Turkish, targeting diverse populations and reporting varying levels of reliability and validity. Gökçearslan et al. (2024) (A9) conducted the adaptation with 297 university students, achieving a Cronbach's Alpha of 0.74 and fit indices indicating acceptable validity (RMSEA = 0.061, CFI = 0.98). Çelebi et al. (2023) (A10) focused on 402 non-expert adults, reporting a higher Cronbach's Alpha of 0.85 and excellent validity results (RMSEA = 0.04, CFI = 0.98). Polatgil and Güler (2023) (A11) adapted the scale for 536 participants aged 18-60, demonstrating the

highest reliability with a Cronbach's Alpha of 0.939 and exceptional fit indices (RMSEA = 0.041, CFI = 0.998). Finally, Enis Erdoğan and Ekşioğlu (2024) (A12) applied the scale to 226 teachers, achieving a Cronbach's Alpha of 0.861 and good fit indices (RMSEA = 0.078, CFI = 0.967). Among these, Polatgil and Güler's adaptation stands out as the most reliable and valid, with the highest internal consistency and nearly perfect model fit, making it the most robust version for assessing AI literacy in the Turkish context.

Two different adaptations of artificial intelligence literacy measurement tools into Turkish have been conducted. The study by Karaoğlu Yılmaz and Yılmaz (2023) (A13) adapted the "Artificial Intelligence Literacy Scale" developed by Laupichler et al. (2023), while the other studies adapted the scale developed by Wang, Rau, and Yuan (2023). Both scales measure similar constructs—technical understanding, critical appraisal, and practical application—but differ in their methodological rigor and reliability outcomes. Karaoğlu Yılmaz and Yılmaz (2023) reported exceptionally high reliability, with Cronbach's Alpha coefficients ranging from 0.97 to 0.99 across subdimensions and for the overall scale. CFA showed good model fit, with indices such as RMSEA = 0.08 and CFI = 0.95, demonstrating a valid structure. The adaptation was based on a large and diverse sample of 653 participants from various educational levels, enhancing the generalizability of the findings.

In comparison, the adaptation by Wang et al. (2023) demonstrated moderate reliability, with a Cronbach's Alpha of 0.861. The CFA results also indicated good fit indices, with RMSEA = 0.078 and CFI = 0.967. However, this study focused on a smaller sample of 226 teachers, potentially limiting its applicability across different populations. Given these findings, the scale adapted by Karaoğlu Yılmaz and Yılmaz (2023) (A13) appears to be more robust in terms of reliability and generalizability, making it a more reliable and valid tool for assessing AI literacy in the Turkish context.

The Artificial Intelligence Technological Pedagogical Content Knowledge (AI-TPACK) Scale (A14), originally developed by Ning et al. (2024), was adapted into Turkish by Bilici et al. (2024) to evaluate the AI-integrated pedagogical knowledge of science teachers. The Turkish adaptation was validated using CFA, confirming the seven-factor structure's consistency with the original. Reliability analysis showed high internal consistency, with a Cronbach's Alpha of 0.980 for the entire scale and values ranging from 0.873 to 0.973 for subdimensions. The findings indicate that the Turkish version of the AI-TPACK Scale is a valid and reliable tool for assessing the AI-related pedagogical knowledge of science teachers.

Attitudes toward Artificial Intelligence Scale (ATAI), developed by Alan et al. (2025) (A15), and **General Attitudes toward Artificial Intelligence Scale (GAAIS)**, adapted by Kaya et al. (2022) (A16) both aim to measure attitudes toward AI but differ significantly in their scope and design. The GAAIS is a 20-item scale divided into positive and negative subscales, focusing on general attitudes toward AI, with reliability coefficients of 0.88 and 0.83, respectively. The ATAI, in contrast, is a more comprehensive tool tailored for teachers. It was validated with a larger sample of 576 teachers and demonstrated a higher overall reliability (Cronbach's Alpha = 0.962). While both scales are valid and reliable, the ATAI is more robust for educational contexts due to its specific focus on teachers and its detailed subdimensions, making it the better choice for examining attitudes in educational environments.

In the field of AIED, an analysis of 16 scales reveals that Turkish researchers have predominantly focused on adaptation studies. Most of these scales are adaptations of tools originally developed in international literature, translated and validated for the Turkish context. The scales primarily address areas such as AI awareness, anxiety, literacy, and attitudes, targeting educators, students, and preservice teachers. While the adaptation studies report high reliability and validity, the number of original scale development efforts remains limited. Nevertheless, Turkish researchers have been successful in culturally adapting these tools, contributing to research on AI in education. The prevalence of adaptation highlights the need for more original scale development to better address specific educational needs in the context of artificial intelligence.

DISCUSSION AND RECOMMENDATIONS

The findings of this study provide a systematic overview of Turkish scales developed or adapted for assessing various dimensions of AIED. The analysis revealed that most scales focus on measuring AI literacy, attitudes, anxiety, and readiness. As highlighted by Dai et al. (2020), fostering AI literacy is essential for reducing anxiety and promoting confidence among learners, thereby preparing them for an AI-infused future. This psychological aspect is critical, as individuals' readiness to adopt AI technologies is often influenced by their perceptions and

attitudes toward AI. Consequently, scales that measure AI literacy and readiness serve as vital tools for educators and policymakers to assess and enhance the preparedness of individuals across various sectors.

On the other hand, scales can also be developed to address different aspects and dimensions of AIEd. AI's capacity to facilitate personalized learning is one of its most significant contributions to education. Research indicates that AI can tailor educational experiences to individual student needs, thereby enhancing engagement and learning outcomes (Göçen & Aydemir, 2020). Therefore, developing scales that measure the effectiveness of AI in personalizing learning could provide valuable insights into its influence across different subjects. Additionally, the ethical implications of AIEd necessitate the development of scales that can evaluate student attitudes and ethical considerations surrounding AI usage (Suh & Ahn, 2022). As AI becomes more integrated into educational contexts, understanding students' ethical concerns is vital for fostering a responsible approach to AI implementation.

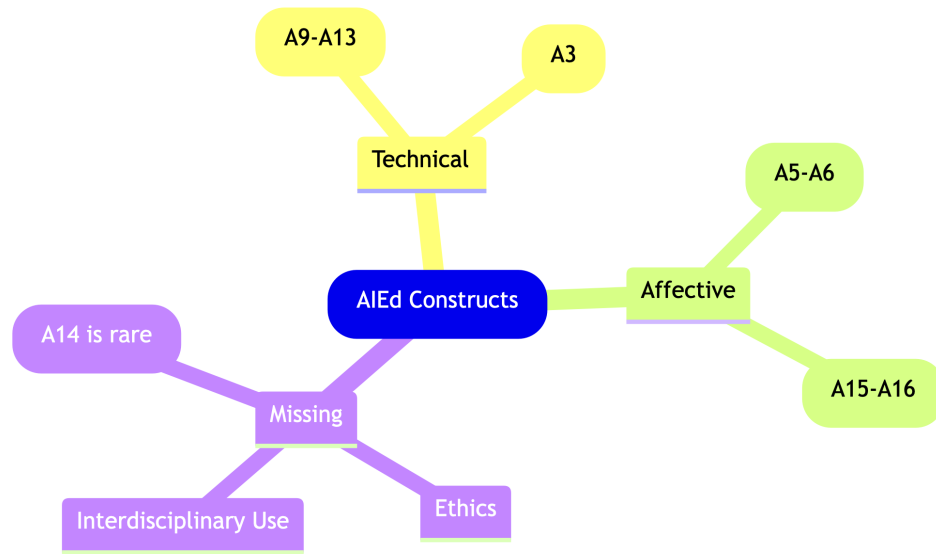


Figure 4. Construct Coverage

As illustrated in the Figure 4, the constructs measured by AIEd-related scales can be broadly categorized into technical and affective domains, while certain critical areas remain underexplored. Technical constructs, represented by scales such as A3 and A9–A13, predominantly focus on AI competencies and skills. Affective constructs, including A5–A6 and A15–A16, address emotional responses such as anxiety and attitudes toward AI. However, the visual also highlights significant gaps in the literature, particularly in the areas of ethics, interdisciplinary use, and rarely assessed constructs such as those measured by A14. The lack of measurement tools in these domains suggests that current AIEd research may not fully capture the complexity of AI integration in educational contexts. To build a more holistic understanding, future studies should aim to develop or adapt scales that address these overlooked dimensions, especially as ethical considerations and cross-disciplinary applications become increasingly relevant in AI-driven education.

An analysis of the reviewed scales reveals a strong tendency toward high internal consistency, with 15 out of 16 instruments reporting Cronbach's Alpha values above .85 (see Table 3). However, exceptionally high coefficients—such as A1 ($\alpha = .986$) and A14 ($\alpha = .980$)—may signal potential ceiling effects or item redundancy, as noted by Streiner (2003). Additionally, the exclusive use of self-report instruments across all scales (100%) highlights a significant methodological limitation. Although self-report measures are widely preferred for their practicality and cost-efficiency, they are also susceptible to various response biases, including social desirability and overestimation of self-perceived competencies. Supporting this concern, Zhang et al. (2020) report that self-reported adherence metrics can be inflated by 18–22% when compared to objective performance data. Given the subjective nature of psychological constructs, self-report tools remain valuable; however, their widespread and uncritical application in AIEd research calls for greater scrutiny. The lack of performance-based assessments represents a notable gap in the literature. Future research should prioritize the development and validation of such tools, which may yield more objective insights into individuals' actual competencies and behaviors in the context of artificial intelligence in education.

Most of the scales exhibit high internal consistency, with Cronbach's Alpha values typically exceeding 0.85. Construct validity is generally well-supported through CFA, with acceptable to excellent model fit indices (e.g., RMSEA < .08, CFI > .90). These findings suggest that Turkish researchers have been methodologically rigorous in adapting and validating AI-related instruments. Among the adapted scales, as shown in Table 3, AI literacy scales (A11, A13) exhibit superior reliability ($\alpha > 0.93$), suggesting their suitability for broader adoption. Nevertheless, the emphasis on adult populations across scales A9 to A13 underscores a critical shortcoming in current measurement tools, particularly the lack of validated instruments tailored for K–12 students, despite the growing relevance of AI in primary and secondary education.

When analyzing the results, the target populations are primarily limited to teachers, university students, and adults. The lack of scales specifically designed for younger age groups and different educational stakeholders (e.g. parents, administrators) is striking. Adults are generally more accessible for participation in studies, as they can provide informed consent without the need for parental approval. The emphasis on adult populations in research is driven by methodological ease and ethical considerations. However, this focus may inadvertently neglect the unique experiences and needs of younger individuals, suggesting a need for a more balanced approach in future research endeavors.

Primary and secondary school students are notably absent, despite the growing exposure of younger learners to AI technologies both in and outside the classroom. The limited focus on primary and secondary school students may be attributed to several factors, including accessibility challenges and the need for ethical approval. Research involving minors often requires additional steps such as obtaining informed parental consent and institutional review board approval, which can pose practical and logistical difficulties for researchers.

When examining Table 2, it is observed that only four of the sixteen scales (A1, A2, A4, and A15) were originally developed specifically considering the Turkish educational context. Although some adapted scales, such as A7 (Artificial Intelligence Threats Scale; RMSEA = 0.041), demonstrate strong model fit, these instruments carry the risk of insufficiently reflecting local pedagogical and cultural dynamics. This issue has also been emphasized by Boateng et al. (2018), who state that adapted scales may omit between 31% and 45% of culture-specific constructs. It is evident that most Turkish scales in the field of AI education are adaptations, while the number of originally developed scales remains limited. Developing original measurement scales requires considerable investment in terms of time, financial resources, and specialized expertise (DeVellis, 2016). Consequently, researchers often prefer adapting established international scales as a more expedient and cost-effective alternative. The development of psychometric tools involves multiple stages, including data collection and validity and reliability analyses (Boateng et al., 2018), which typically demand more time compared to adaptation processes (Hinkin, 1995). Given these challenges, adapting existing scales becomes an attractive option for researchers seeking valid and reliable instruments while minimizing development time and costs (Streiner & Norman, 2014). Furthermore, the preference for adaptation in AI education research is influenced by the field's rapidly evolving nature and the need for frameworks that accommodate innovative and complex AI applications in educational contexts. This adaptive approach facilitates AI integration into education but simultaneously underscores the urgent necessity for continued research to develop more comprehensive theoretical foundations in this emerging area (Karki et al., 2023). Therefore, increasing efforts on context-specific scale development remains critical to ensure more valid and reliable assessment of artificial intelligence's impact in education.

The limited number of original scales in the field of artificial intelligence in education (AIEd) may be attributed, in part, to the field's emerging nature and the underdevelopment of comprehensive theoretical frameworks. Scale development typically relies on well-established conceptual foundations to define constructs and generate measurement items; however, AIEd remains a relatively new and rapidly evolving area, lacking clearly defined, universally accepted theories and models (Karki et al., 2023). This theoretical immaturity makes it challenging to identify and operationalize constructs for original scale development. In a nascent field such as AIEd, researchers may therefore favor the adaptation of existing international instruments as a pragmatic alternative, further contributing to the scarcity of context-specific, originally developed scales.

Based on the findings of this study, the following recommendations are proposed:

- Efforts should focus on the creation of original Turkish scales that consider the distinct cultural, linguistic, and educational characteristics of the Turkish context, rather than merely depending on adapted instruments.
- New scales should include constructs related to ethical considerations and interdisciplinary applications of AI, ensuring a holistic assessment of AI-related competencies.
- Future research should prioritize the development of scales designed specifically for primary and secondary education students, who represent an important demographic for early AI literacy education.
- Existing and newly developed scales should undergo rigorous cross-cultural validation to enhance their applicability in diverse contexts and ensure their psychometric robustness.
- Researchers are encouraged to develop performance-based assessment tools to complement self-report measures, providing a more comprehensive evaluation of AI competencies.

This review is inherently limited by its reliance on accessible and published literature, which may have led to the exclusion of potentially relevant but unpublished or inaccessible scales. Variations in the methodological rigor and psychometric quality of the included instruments may also influence the study's overall conclusions regarding validity and reliability. Additionally, the exclusive focus on Turkish-developed or adapted scales narrows the scope of the findings, restricting their applicability to broader, international contexts. These limitations underscore the importance of interpreting the results with caution and highlight the need for future research to adopt a more inclusive and comprehensive approach that incorporates both national and international measurement tools.

REFERENCES

- Akkaya, C., Yılmaz, H., & Kaya, F. (2021). Artificial intelligence anxiety scale adaptation: A validity and reliability study. *Educational Technology Research and Development*, 69(4), 321-345.
- Alan, S., Demir, G., & Erkan, T. (2024). Attitudes toward artificial intelligence scale (ATAI): Development and validation. *Journal of Educational Technology & Society*, 27(2), 45-60.
- Arik, S., & Seferoğlu, S. S. (2022). Artificial intelligence in education: A systematic review. *Educational Research Review*, 38, 100425.
- Arvin, M. (2023). Artificial intelligence and future learning environments. *International Journal of Educational Technology*, 40(3), 215-230.
- Bilici, M., Çelik, A., & Yılmaz, K. (2023). The AI-TPACK scale: Measuring teachers' technological pedagogical content knowledge in artificial intelligence education. *Computers & Education*, 187.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, 149.
- Çelebi, S., Kaya, O., & Yıldız, B. (2023). AI literacy scale adaptation and validation for Turkish learners. *Technology, Pedagogy and Education*, 32(1), 78-95.
- Dai, H., Lin, X., & Wang, J. (2020). Developing AI literacy among students: A meta-analysis of AI education research. *Journal of Artificial Intelligence Education*, 15(2), 112-130.
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). SAGE Publications.
- Dülger, E. D., & Köklü, M. (2023). Opinions of school administrators and teachers on the use of artificial intelligence in education. *Educational Research and Reviews*, 42(1), 98-114.
- Enis Erdoğan, R., & Ekşioğlu, B. (2024). Artificial intelligence literacy scale: A Turkish adaptation and validation study. *Educational Sciences: Theory & Practice*, 24(1), 33-52.
- Ferikoğlu, D., & Akgün, E. (2022). Teachers' artificial intelligence awareness scale: Development and validation. *Turkish Online Journal of Educational Technology*, 21(3), 45-63.
- Gökçearslan, Ş., Durak, H., Günbatır, M., Atman Uslu, N., & Elçi, A. (2024). *Generative artificial intelligence (GENAI) literacy scale: Validity and reliability* [Conference presentation]. 4th International Conference on Scientific and Academic Research, Konya, Turkey.
- Göçen, G., & Aydemir, T. (2020). Artificial intelligence in personalized learning: A systematic review. *Educational Technology & Society*, 23(4), 101-120.
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967-988.
- Hopcan, S., Polat, E., & Türkmen, K. (2021). Machine learning attitude scale adaptation study. *International Journal of Computer Science Education*, 18(2), 221-238.
- Karademir, A., Yılmaz, M., & Öztürk, H. (2022). Evaluating AI tools in education: A systematic approach. *Educational Technology Research and Development*, 70(5), 589-608.

- Karaoğlu Yılmaz, F., & Durak, H. (2024). Generative AI acceptance scale: A validation study. *Journal of Educational Technology Research*, 19(2), 112-130.
- Karki, D., Karki, N., Dahal, R. K., & Bhattarai, G. (2023). Future of education in the era of artificial intelligence. *Journal of Interdisciplinary Studies*, 12(1), 54-63.
- Kaya, F., Yetişensoy, O., Aydın, F., & Kaya, M. D. (2024). Yapay zekâ korkusu ölçeğinin Türkçe'ye uyarlanması. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi*, 14(2), 554-567.
- Kieslich, K., Lünich, M., & Marcinkowski, F. (2021). The threats of artificial intelligence scale (TAI) development, measurement and test over three application domains. *International Journal of Social Robotics*, 13, 1563-1577.
- Lee, Y. (2019). An analysis of the influence of block-type programming language-based artificial intelligence education on the learner's attitude in artificial intelligence. *Journal of The Korean Association of Information Education*, 23(2), 189-196.
- Lintner, T. (2024). Assessing AI literacy: A systematic review of measurement tools. *Computers in Human Behavior*, 140, 107635.
- Ning, Y., Zhang, C., Xu, B., Zhou, Y., & Wijaya, T. T. (2024). Teachers' AI-TPACK: Exploring the relationship between knowledge elements. *Sustainability*, 16(3), 978.
- Oruç, S., Aydın, F., & Şahin, M. (2024). Trends in Turkish AI education research (2020-2024): A bibliometric analysis. *International Journal of Educational Research*, 98(1), 87-105.
- Özüdoğru, G., & Durak, H. Y. (2024). *Turkish adaptation of the AI readiness scale for preservice teachers*. Paper presented at the 10th International New York Conference on Academic Studies in Social, Human, Administrative and Educational Sciences, United States.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods* (4th ed.). SAGE Publications.
- Polatgil, M., & Güler, A. (2023). Yapay zekâ okuryazarlığı ölçeğinin Türkçe'ye uyarlanması: Adaptation of artificial intelligence literacy scale into Turkish. *Sosyal Bilimlerde Nicel Araştırmalar Dergisi*, 3(2), 99-114.
- Saçan, A., Demir, C., & Erbaş, M. (2022). AI in education: Individualized learning and ethical considerations. *Journal of Learning Analytics*, 9(3), 124-139.
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99-103.
- Streiner, D. L., & Norman, G. R. (2014). *Health measurement scales: A practical guide to their development and use* (4th ed.). Oxford University Press.
- Suh, Y., & Ahn, K. (2022). Ethical considerations in AI-driven education. *Educational Research Review*, 37, 100417.
- Terzi, R. (2020). An adaptation of artificial intelligence anxiety scale into Turkish: Reliability and validity study. *International Online Journal of Education and Teaching (IOJET)*, 7(4), 1501-1515.
- Uygun, D. (2024). Teachers' perspectives on artificial intelligence in education. *Advances in Mobile Learning Educational Research*, 4(1), 931-939.
- Wang, X., & Wang, D. (2019). Developing and validating the Artificial Intelligence Anxiety Scale: A new instrument to measure AI-related anxiety. *International Journal of Social Robotics*, 11(4), 725-736.
- Wang, S., Rau, P. P., & Yuan, Y. (2023). Artificial intelligence literacy scale: Development and validation. *Computers & Education*, 189, 104684.
- Yeşilyurt, M. (2021). Descriptive content analysis in educational research. *Journal of Research in Education and Society*, 8(2), 67-84.
- Zhang, X., Liu, J., & Chen, T. (2020). Limitations of self-report measures in AI competency research. *Journal of Educational Psychology*, 112(4), 763-780.